



Top administration received the August 1978 Report favorably and expected to bring the matter to the Board of Regents in their October 1978 meeting. However, the Board did not address the issue as planned, and in April of 1979, Vice-President Carter requested and received three alternate equipment plans, each with three alternate financing, budget, and funding options:

- Option 1. Proceed with full acquisition of equipment recommended in Computer Task Force Report.
- Option 2. Proceed with acquisition of a dual processor IBM-compatible computer system to be shared by academic and administrative users. Additional equipment to be acquired in a later procurement.
- Option 3. Proceed with minimal acquisition of IBM-compatible computer system to be shared by academic and administrative users. To be followed by an immediate procurement to offset rapid overload.

In June of 1979, Vice-President Carter requested and received a Computing Center budget based on Option 3. This budget was approved in late August, 1979, and the Computing Center staff proceeded to prepare state justification and specifications documents which resulted in the current acquisition.

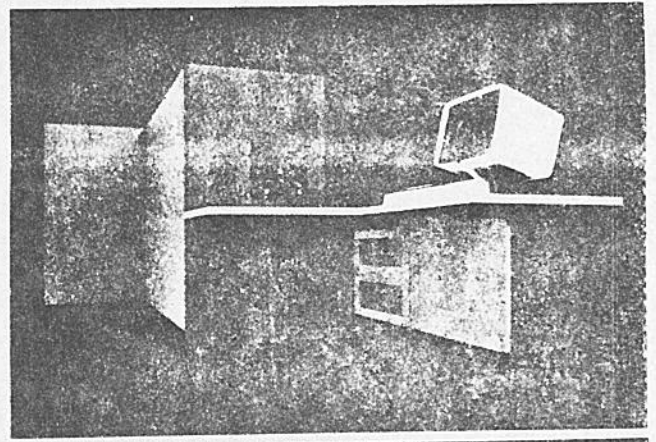
It is generally agreed that the AS/5000 will meet less than half the needs projected through 1983 by the August 1978 Computer Task Force Report. All subsequent budget and justification documents have clearly stated that this procurement only partially meets short term requirements. It is a very practical first step within current operating funds limitations, but another acquisition process should be initiated as soon as possible. Space commitments are also urgently needed.

Please consult the following brochure supplied by National Advanced Systems for more information concerning the AS/5000.



| Machine Type  | Average Relative Performance | Machine Type | Average Relative Performance | Machine Type  | Average Relative Performance | Machine Type    | Average Relative Performance |
|---------------|------------------------------|--------------|------------------------------|---------------|------------------------------|-----------------|------------------------------|
| IBM/4331      | .15                          | NAS/AS-3000  | 1.00                         | NAS/AS-7000-N | 2.40                         | AMD/470/V7A     | 4.80                         |
| IBM/3135      | .20                          | IBM/3158-3   | 1.00                         | IBM/3168-3    | 2.75                         | NAS/AS-7000-DPC | 5.00                         |
| IBM/3138      | .25                          | IBM/3031     | 1.20                         | IBM/3032      | 2.75                         | IBM/3033        | 5.80                         |
| IBM/3145      | .35                          | NAS/AS-5000  | 1.35                         | AMD/470/V5    | 2.75                         | AMD/470/V7      | 6.00                         |
| IBM/3148      | .45                          | IBM/3158-AP3 | 1.45                         | NAS/AS-7000   | 3.15                         | AMD/470/V8      | 7.40                         |
| IBM/3155-II   | .70                          | IBM/3031-AP  | 1.75                         | AMD/470/V7B   | 3.40                         | IBM/3033-AP     | 8.80                         |
| IBM/4341      | .75                          | IBM/3165-II  | 2.00                         | IBM/3033-N    | 3.60                         | IBM/3033-MP     | 9.80                         |
| NAS/AS-3000-N | .75                          | IBM/3165-III | 2.20                         | IBM/3168-AP3  | 3.90                         |                 |                              |
| IBM/3158-1    | .90                          | IBM/3168-1   | 2.40                         | AMD/470/V6-II | 4.40                         |                 |                              |





## **National Advanced Systems**

# AS/5000

### Product Overview

Employing advanced technology unsurpassed by its IBM counterparts, National Advanced Systems' AS/5000 is designed to deliver performance superior to both the IBM 370/158-3 and 3031 at a significantly lower cost.

The AS/5000's faster cycle time, larger buffer store and increased Reloadable Control Storage capacity, together with other performance enhancing features, make it a natural upgrade path for 148,158 or AS/3000 users.

In addition, the AS/5000 takes up less floor space, requires approximately 50% less power and generates about half as much heat as comparable IBM systems.

### Full IBM Compatibility

To help protect your investment in people, programming and training, National's AS/5000 is completely operationally compatible with the IBM 370/158. All peripheral devices designed to interface to a 158 will run on the Advanced System/5000. Additionally, all user software, except time-dependent code, developed for an IBM 370/158 may be used with the AS/5000. Operators, application programmers and systems programmers trained on the 158 can operate and program the Advanced System/5000 immediately, without further training.

### Superior Technology Means Superior Performance

#### 1) Reloadable Control Storage (RCS)

With 16K 72-bit words, National's Advanced System/5000 offers twice as much Reloadable Control Storage for microcode as IBM's 3031. As a result, the AS/5000 has the capability for more firmware assist features, including the standard Virtual Machine Assist (VMA) and Extended Control Program Support (ECPS), plus optional features like VM ECPS including Control Program Assist (CPA), Extended Virtual Machine Assist (EVMA), and Virtual Interval Timer Assist (VITA), which together can provide a performance improvement for VM of 10% to 35%. In addition, both 14XX and 70XX emulation, as well as OS/DOS emulation are standard with the AS/5000.

#### 2) Spare Channel

To provide configuration flexibility, the AS/5000 offers seven multiplexer channels to be configured as one byte, five block and a spare byte, or two byte, four block and a spare block.

### Microprogramming Overview

The hardware operation of our Advanced System/5000 large scale computer system is controlled by microcode which resides in Reloadable Control Storage (RCS). 16K words of RCS are provided on the AS/5000, for which only 8K words are required for compatibility with the System/370 architecture.

Additional firmware is available to provide support for certain virtual storage operating systems. The use of these firmware assist features in conjunction with the relevant operating systems can result in a significant reduction in operating system overhead and a consequent improvement in performance.

### AS/5000 Firmware Options

The Advanced System/5000 offers five firmware options:

- ☐ Emulators—14XX Emulator, 70XX Emulator, OS/DOS Compatibility Feature
- ☐ Virtual Machine Assist (VMA)
- ☐ Extended Control Program Support (ECPS)—VS/1 Assist, Virtual Machine Assist (VMA)
- ☒ VM ECPS—Control Program Assist (CPA), Extended Virtual Machine Assist (EVMA), and Virtual Interval Timer Assist (VITA)
- ☐ System/370 Extended Facility (S/370 EF)

The AS/5000 retains all firmware features available on the 370/158-3 (emulators, VMA

and ECPS), plus it provides S/370 Extended Facility and VM ECPS firmware enhancements. The AS/5000 offers more firmware flexibility than any other machine in its class. And, with 16K words of RCS, its microcode capability is expandable to provide future compatibility.

### VMA Enhances AS/5000 Performance

Virtual Machine Assist (VMA) firmware increases system performance by dramatically reducing VM/370 overhead associated with privileged instruction simulation. VMA emulates those privileged instructions issued by a virtual machine which are associated with PSW handling, paging, storage keys and I/O. The use of VMA results in a VM/370 performance improvement on the AS/5000 in the range of 10% to 50%.

### ECPS Increases VS/1 Performance

Extended Control Program Support (ECPS) contains both Virtual Machine Assist (VMA) for VM/370 and a VS/1 Assist for the VS/1 operating system. Both assists can be used concurrently as in a VS/1 operating system running under VM/370, or separately. For VM/370, VMA can provide a 10% to 50% performance improvement. And for VS/1, performance gains range from 5% to 14%.

The VS/1 Assist included in the Extended Control

Program Support (ECPS) increases internal CPU performance by reducing VS/1 operating system overhead. The reduction in overhead is achieved by moving certain heavily used operating system software functions, such as page fix, dispatch and trace, into microcode. The VS/1 routines moved into microcode are related to memory management, paging, interrupt handling and trace.

### VM ECPS Extends VM/370 Performance

VM Extended Control Program Support contains Control Program Assist (CPA), Extended Virtual Machine Assist (EVMA), and Virtual Interval Timer Assist (VITA). Working in conjunction with VMA contained in the Extended Control Program Support (ECPS), the VS/1 Assist in ECPS is also available for use with a VS/1 system running under VM/370.

CPA increases VM/370 systems' internal CPU performance by reducing VM/370 operating system overhead. The reduction in overhead is achieved by moving certain heavily used VM/370 system software functions, such as COW translation, maintenance of dynamic storage and dispatching, into microcode.

The Extended Virtual Machine Assist (EVMA) extends the microcode support for privileged operation simulation. The support for the PSW handling instructions in



EVMA is extended beyond the support given them by VMA.

Virtual Interval Timer Assist (VITA) reduces VM overhead and improves timing accuracy for operating systems running under VM. With VITA the virtual machine interval timer is updated via microcode. Additional microcode assist is given for TLB and I/O privileged instructions. Benchmarks of EVMA have shown a performance improvement of up to 35% over that achieved by VMA alone.

#### **System/370 EF Improves Performance**

System/370 Extended Facility supports IBM's MVS/ System Extensions program product (5740-XE1). MVS/SE in conjunction with S/370 EF provides new functions and can result in performance improvements ranging between 12% and 15%.

#### **New Functions:**

- ☐ Additional main memory protection for the first 512 bytes enhances operating system integrity.
- ☐ Common segment bit improves performance by sharing segments of Virtual Storage common to multiple address spaces.
- ☐ New instructions enhance performance by providing:
  - SVC Assist
  - Trace Assist
  - Dispatching
  - I/O Interrupt Assist
  - Program Interrupt Assist
  - Selective Purge of TLB

#### **National Microcode Means Performance**

Our firmware enhancements assist the Advanced System/5000 in delivering greater price/performance efficiency than any other machine in its class. In conjunction with the relevant

operating system, each firmware option provides performance characteristics needed to increase internal CPU performance dramatically.\*

#### **Full Service Support**

Our commitment to delivering superior performance and reliability at a very low cost extends beyond the design and development of our equipment. We back the AS/5000 and all our computer products with full service support. From our worldwide network of expert Field Engineers and System Support Representatives, to our wide range of program products, to our virtually limitless number of financing programs, National offers total capabilities to help protect your computer investment—now, and in the future.

| Firmware Option                                          | Operating System | Performance Improvement Range |
|----------------------------------------------------------|------------------|-------------------------------|
| S/370 Extended Facility                                  | MVS/SE           | 12% to 15%                    |
| Virtual Machine Assist                                   | VM/370           | 10% to 50%                    |
| VM Extended Control Program Support (in addition to VMA) | VM/370           | 10% to 35%                    |
| Extended Control Program Support                         | VS/1             | 5% to 14%                     |

\*The performance improvement realized by each of the firmware options depends upon the customer environment and the use made of the facilities provided. The performance figures stated, however, represent a range of improvement that can be expected in the majority of installations. No assurance can be given that an individual user will achieve the same results.

**Specifications****Virtual Storage Capability****Principal Operating Systems****AS/5000**

Standard  
MVS  
MVS-SE  
VM/370  
VM/370-SE  
SVS  
VS/1  
DOS/VS  
DOS/VSE (370 mode)

**Number of CPUs****Processor Cycle Time****Channels****Transfer Rates**

Block  
Byte

**Aggregate Data Rate****Memory**

Minimum Size  
Maximum Size

**Increment****Read Cycle** (16 Bytes)**Write Cycle** (8 Bytes)**Bytes/Cycle****Cache Buffer**

Capacity  
Access Time  
Bytes/Cycle

1

92 ns

6

1.5 MB/sec

100 KB/sec

6.75 MB/sec

16K NMOS

ECC std.

2 MB std.

8 MB

1 MB

460 ns

460 ns

8

32K

184 ns

8

**RCS**

Assist Features

**Operating Environment**

Temperature  
Humidity

**Non-operation**

Temperature  
Humidity

**Power Requirements****Phase/Wire****Power Consumption****Heat Output****Space Requirements****Weight****Standard Features**

Console CRT  
370/EF  
OS/VS1 ECPS  
VMA

**Optional Features**

Additional Memory  
Direct Control  
Console Printer  
Remote CRT  
VM ECPS

16K (72 bit) Words  
370/EF  
OS/VS1 ECPS  
VMA  
VM ECPS

60° to 90°F (16° to 32°C)  
20% to 80% RH  
(Max. wet bulb 78°F or  
26°C)

50° to 110°F (10° to 43°C)  
8% to 80% RH  
(Max. wet bulb 80°F or  
27°C)

200/208/230 VAC <sup>-20%</sup>/<sub>+10%</sub>  
@ 60HZ

3 Phase/4 Wire

11.4 KVA (Maximum)

34,500 BTU/hr (8,694 kcal/hr)

50.5 ft.<sup>2</sup> (4.69m<sup>2</sup>)

4,090 lbs. (1,855 kg)

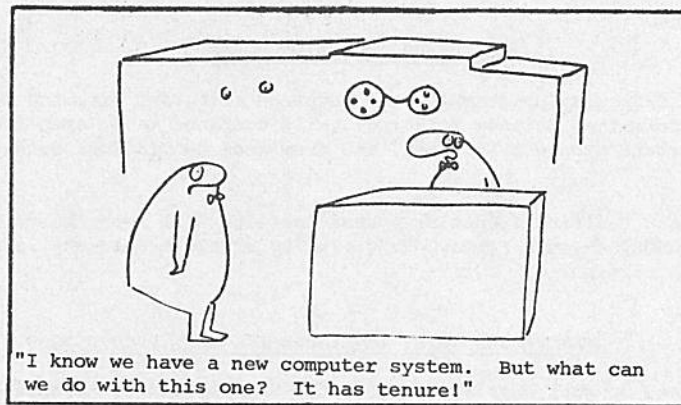


3145 Porter Drive  
Palo Alto, CA 94304  
Telephone (415) 856-5000



# COMPUTER HUMOR

Contributions to "COMPUTER HUMOR" are always welcomed.



## VM/370

The installation of the AS/5000 brings with it a new concept: VM/370 (Virtual Machine Facility/370). VM is an operating system that manages the resources of a single IBM System/370 compatible computer so that multiple computing systems (virtual machines) appear to exist. For example, VM could be used to implement separate "virtual" academic and administrative machines with shared hardware, software, technical and operations support.

We are currently operating our old MVT operating system under the control of VM. A new MVT is being implemented with support for the new disk drives and remote terminals. The MUSIC (McGill University System for Interactive Computing) timesharing operating system, developed by McGill University and distributed by IBM, will also be implemented under VM in July, and should be in general use by the start of the Fall semester.

A "popular" saying with regard to the concept of a virtual system may prove useful in understanding it:

- If you have it, and know you have it, it's REAL.
- If you have it, and don't know you have it, it's TRANSPARENT.
- If you don't have it, but think you have it, it's VIRTUAL.

## MUSIC

Current plans call for the implementation of MUSIC rather than IBM's TSO or CMS timesharing software. As stated previously, MUSIC is the McGill University System for Interactive Computing. It is a high performance operating system that comes complete with a comprehensive package of programs and facilities designed to meet the varied needs of interactive computing. For example, users may concurrently be performing such diverse activities as problem solving, program development, file editing, word processing, and batch processing.

MUSIC will be discussed in greater detail in a forthcoming issue of BENCHMARKS.

CHANGE IN CHARGE RATES EXPECTED

All Computing Center charge rates will be reevaluated this summer. The rate for CPU time will be increased to 21¢ per second to reflect the increase in computing power until new rates are established. This should result in a slightly lower per job cost during the interim.

SUMMER COMPUTING HOURS

The computer room, student/faculty programming and keypunch area, and dispatch desk, which are all in the Computing Center (Information Science Building), will continue to be open twenty-four hours a day from 7:30 a.m. Monday through midnight Saturday, and from noon to midnight Sunday, when school is in session.

The Business Administration Building Computing Access Facility will open from 7:30 a.m. to 4:00 a.m. the following morning, Monday through Friday, 7:30 a.m. to midnight Saturday, and noon to midnight Sunday, when school is in session.

REMOTE JOB ENTRY (RJE) OPERATIONS: SECURE JOBS

RJE operations are designed to work best if users of the RJE facilities in BA152 enter (read) their own jobs into the central computer and retrieve their own program listings from the self service output boxes. Users should request secured output only when absolutely necessary.

The improved turnaround and corresponding increase in jobs that will be processed by the AS/5000 this summer will place a severe load on the dispatch assistants that file the printouts for retrieval. Requests for secured output adds another step to the distribution process resulting in increased wait time for all users.

Hopefully, the need for secured output will diminish with improved turnaround. Therefore, all faculty members are urged to discourage use of secured output by their students.

\* \* \* \* \*

\*                      USAGE      STATISTICS      FOR      SPRING      1980                      \*

\* \* \* \* \*

IBM/360 JOB STATISTICS BY JOB CLASS

| JOB CLASS | JANUARY |         | FEBRUARY |         | MARCH  |         | APRIL  |         |
|-----------|---------|---------|----------|---------|--------|---------|--------|---------|
|           | NUMBER  | PERCENT | NUMBER   | PERCENT | NUMBER | PERCENT | NUMBER | PERCENT |
| A         | 5,246   | 45.4    | 10,807   | 39.3    | 13,047 | 42.6    | 16,636 | 44.2    |
| B         | 1,026   | 8.9     | 1,789    | 6.5     | 1,339  | 4.4     | 1,817  | 4.8     |
| C/D/G/I   | 1,565   | 13.5    | 2,426    | 8.8     | 2,416  | 7.9     | 2,982  | 7.9     |
| J/K/M/N   | 431     | 3.7     | 363      | 1.3     | 297    | 1.0     | 360    | 1.0     |
| 1         | 900     | 7.8     | 2,856    | 10.4    | 2,350  | 7.7     | 2,423  | 6.4     |
| 2         | 2,062   | 17.9    | 6,201    | 22.5    | 3,838  | 12.5    | 5,798  | 15.4    |
| 3         | 0       | 0       | 485      | 1.8     | 1,105  | 3.6     | 399    | 1.1     |
| 4         | 320     | 2.8     | 2,606    | 9.5     | 6,216  | 20.3    | 7,255  | 19.3    |
| TOTAL     | 11,550  | 100.0   | 27,533   | 100.0   | 30,608 | 100.0   | 37,670 | 100.0   |

IBM/360 TURNAROUND STATISTICS

| AVERAGE TIME (MINUTES) |         |          |       |        |
|------------------------|---------|----------|-------|--------|
| JOB SPENT              | JANUARY | FEBRUARY | MARCH | APRIL  |
| Waiting To Execute     | 13.82   | 18.40    | 43.99 | 113.77 |
| Executing              | 3.41    | 2.76     | 3.53  | 4.24   |
| Waiting to Print       | 12.48   | 8.98     | 11.41 | 16.19  |
| Printing               | 2.65    | 1.42     | 1.56  | 1.48   |
| Read In To Print Time  | 32.82   | 31.84    | 60.94 | 136.15 |



IBM/360 USAGE BY MAJOR AREAS  
1/1/80 to 4/30/80

| AREA           | JOBS           |         | LINES             |         | CPU CYCLES (SECS) |         |
|----------------|----------------|---------|-------------------|---------|-------------------|---------|
|                | NUMBER         | PERCENT | NUMBER            | PERCENT | NUMBER            | PERCENT |
| INSTRUCTION    | 93,931         | 87.9    | 37,056,396        | 53.9    | 1,545,979         | 52.5    |
| (ACIS)         | (43,990)       | (41.2)  | (19,397,025)      | (28.2)  | (755,249)         | (25.7)  |
| (CSCI)         | (41,244)       | (38.6)  | (11,438,492)      | (16.6)  | (404,414)         | (13.7)  |
| (OTHER)        | ( 8,697)       | ( 8.1)  | ( 6,220,879)      | ( 9.1)  | (386,316)         | (13.1)  |
| RESEARCH       | 2,417          | 2.3     | 2,591,677         | 3.8     | 512,338           | 17.4    |
| SYSTEMS        | 4,203          | 3.9     | 2,967,513         | 4.3     | 156,002           | 5.3     |
| ADMINISTRATIVE | 6,320          | 5.9     | 26,135,989        | 38.0    | 730,201           | 24.8    |
| TOTAL          | <u>106,871</u> | 100.0   | <u>68,751,575</u> | 100.0   | <u>2,944,520</u>  | 100.0   |

| AREA           | CPU PROC (HRS) |         | EXCP'S            |         | CHARGES (\$)        |         |
|----------------|----------------|---------|-------------------|---------|---------------------|---------|
|                | NUMBER         | PERCENT | NUMBER            | PERCENT | AMOUNT              | PERCENT |
| INSTRUCTION    | 566            | 99.6    | 45,908,682        | 60.4    | 233,621.42          | 58.1    |
| (ACIS)         | (273)          | (48.1)  | (30,234,577)      | (39.8)  | (110,653.00)        | (27.5)  |
| (CSCI)         | (274)          | (48.2)  | (10,600,358)      | (13.9)  | ( 75,438.25)        | (18.8)  |
| (OTHER)        | ( 19)          | ( 3.3)  | ( 5,073,747)      | ( 6.7)  | ( 47,530.17)        | (11.8)  |
| RESEARCH       | 2              | 0.4     | 2,490,985         | 3.3     | 25,874.69           | 6.4     |
| SYSTEMS        | --             | ---     | 8,613,865         | 11.3    | -----               | ---     |
| ADMINISTRATIVE | --             | ---     | 19,039,600        | 25.0    | 142,403.55          | 35.4    |
| TOTAL          | <u>568</u>     | 100.0   | <u>76,053,132</u> | 100.0   | <u>\$401,899.66</u> | 100.0   |

HP 2000 USAGE

| CATEGORY         | JANUARY | FEBRUARY | MARCH | APRIL |
|------------------|---------|----------|-------|-------|
| HP Connect Hours | 4073    | 4389     | 4261  | 3551  |
| HP Log-Ons       | 6133    | 7601     | 5636  | 6547  |

NOTE: Monthly fluctuations in computer usage are due in part to variations in the number of days school was in session. Below is the number of class days for each month.

|          |         |
|----------|---------|
| JANUARY  | 14 days |
| FEBRUARY | 21 days |
| MARCH    | 16 days |
| APRIL    | 23 days |

NTSU GETS A NEW STATISTICAL SOFTWARE PACKAGE: STATISTICAL ANALYSIS SYSTEM (SAS)

SAS is a computer system for data analysis that reportedly "provides all the tools needed for data analysis"; information storage and retrieval; data modification and programming; report writing; statistical analysis and file handling (SAS Users Guide: 1979 Edition, page '3). Conceptually, SAS is similar to SPSS, and users of SPSS should have little difficulty in executing SAS programs. SAS Conventions are very similar to PL/I programming conventions while SPSS is more closely related to FORTRAN. SAS, like SPSS, must have a "D" job class. To execute SAS, simply insert a // EXEC SAS card directly after your job card.

The Computing Center is currently not able to support SAS at the consultation level, although plans are under way to provide consulting support by the fall semester. At the maintenance and documentation levels, however, the Computing Center currently provides full support (see "Software - What We Have And What We Support" in the January issue of BENCHMARKS for a more detailed description of the support levels.)

Potential users of SAS should consult the SAS Users Guide mentioned above and the SAS Introductory Guide by Jane T. Helwig. The SAS Users Guide is available at the University Store.

STAT USERS BEWARE

*The following article is adapted from the Statistical Package for the Social Sciences (Second Edition) Norman H. Nie, C. Hadlai Hull, et.al., 1970, McGraw-Hill, (Pages 3-4), and the SPSS Inc. Newsletter, Number 20, February 1980, (Page 10). It can be applied, however, not only to SPSS, but to all the statistical packages available to the users at NTSU: SAS, OSIRIS, NTSU STATPAK, BMD, and any other computer packages that make complex statistical techniques available to the general academic community. Although this article is aimed specifically at the users of the statistical packages, people who rely on Computing Center staff to do their statistical programming should take note also.*

Users of statistical packages should exercise a good deal of caution with regard to the statistical procedures they invoke. While these packages provide the user easy access to a wide range of advanced statistical techniques, they also carry with them the temptation to "fish." That is, the relative ease with which a user can access a high powered statistical procedure leads him/her to attempt statistical analyses that (s)he does not really understand. Such a situation can lead to "Garbage-in, Garbage-out" (GIGO) research, which carried to its extreme could give all research a bad name.

According to the authors of SPSS, statistical packages can be abused in two ways:

1. Ease of access often means overaccess: Since it is so easy to produce large amounts of information on a single pass of a data file, the researcher is tempted to generate mountains of output "... substituting the crudest form of empiricism for the careful interaction of concepts, hypotheses, and data analysis." The authors further state that (a) there is the problem of overproduction of information. "How does one look at 500 crosstabulation tables?", and (b) "If one generates enough information, one is bound by the laws of probability to have some statistically significant findings. A 100 x 100 correlation matrix generated from a file of totally random numbers produces 5,000 unique correlation coefficients. Within this correlation matrix there will likely be five large correlation coefficients significant at the .001 level, fifty at the .01 level, and 250 at the .05 level." The authors suggest that a good rule-of-thumb is to "...request only those tables, coefficients, etc., for which you have some theoretical expectations based upon the hypotheses in your research design."
2. Uninformed use of the available statistical techniques: Statistical procedures accessed through the use of computer packages are basically blind computational algorithms that apply their formulas to whatever data the user enters; they have little ability to distinguish between proper and improper application of their techniques. If, for example, the SPSS subprogram FACTOR is so directed, it will perform a factor analysis on numerically coded nominal data. Since no assumption of ordering or distance is made at the nominal level of measurement, however, the results of performing a factor analysis (which assumes meaningful numerical distances between categories) will almost always produce meaningless results, although this may not be immediately apparent from the output. As a general rule, then, it is advisable that the user never attempt to use a statistical procedure unless (s)he understands both the appropriate procedure for the type of data at hand and the meaning of the statistics produced by the procedure.

It is virtually impossible for the Computing Center staff to become familiar enough with their customers' data to realize whether or not they are violating assumptions when executing requested statistical procedures. Furthermore, Computing Center personnel have neither the time nor the expertise to provide statistical counseling as to what statistics should be applied to which variables. Therefore, the graduate student and/or faculty member who anticipates utilizing Computing Center staff to do his/her statistical programming should have a clear idea of the variables they want analyzed and the statistical procedure(s) they want used to accomplish this goal.



```

* * * * *
*               S P S S               *
* * * * *

```

Since the last issue of BENCHMARKS which announced the installation of SPSS Release 8.0, Release 8.1 has been installed. This version is essentially the same as Release 8.0, minus most of the "bugs" and with a few modifications. The following is a list of new or modified SPSS conventions, functions, and procedures. Following each keyword descriptor is a brief statement of changes and/or additions that have occurred. These changes/additions in Release 8.1 are not documented in the SPSS Update manual. For a complete list of changes and/or additions to SPSS as of Release 8.0, consult the SPSS Update manual for Releases 7 and 8 on the BENCHMARKS SPSS supplement.

#### NEW CONVENTIONS, FUNCTIONS, AND PROCEDURES

1. ANOVA: With Release 8.1, ANOVA will now compute sums of square and mean squares for one case per cell designs in the absence of Option 3.
2. DATA DEFINITION AND TRANSFORMATION CONTROL CARDS: The limit of 250 symbols on [\*]SELECT IF,MISSING VALUES,ASSIGN MISSING, and PRINT FORMATS cards no longer applies in Release 8.1. The limit is not related to the amount of workspace available for storage of the symbols as they are being parsed.
3. FACTOR: NFACTORS and MINEIGEN on Release 8.1 can now be used to control the number of factors extracted with image factoring.
4. FISHERS EXACT TEST: With Release 8.1, SPSS prints a two-tailed test of significance for this statistic.
5. Keyword THRU: On Release 8.1, the keyword TO may now be used in place of the keyword THRU wherever THRU is called for.
6. MULT RESPONSE: With Release 8.1, MULT RESPONSE will produce tables of groups with groups with pair variables between the groups, rather than pool all variables.
7. SEED: With Release 8.1, the first random number drawn is no longer periodically related to the time of day.

#### KNOWN ERRORS IN RELEASE 8.1

The following is a list of errors known to exist in SPSS Release 8.1. This list is being published so that users may be alerted to possible discrepancies in their output.

##### BREAKDOWN:

- (a) Integer mode, CROSSBREAK FACILITY ONLY. The Chi-square warning message for cells with expected frequencies less than 5 is unreliable. It is not always printed when it should be, and the counts and minimum values listed are not always accurate.
- (b) Option 2 (Exclude cases with missing values on the dependent variable) does not work for CROSSBREAK. Cases with missing values on the dependent variable are included in the calculation of the statistics.
- (c) CROSSBREAK. There is a maximum of 200 non-empty rows and 200 non-empty columns allowed. If these maxima are exceeded, a non-fatal message is printed and processing continues. The manuals do not document this limitation.

##### COMPUTE and IF:

An expression of the form  $Y - 2 ** Z$  is parsed incorrectly as  $Y + (-2) ** Z$ . The result is correct if Z is odd, but incorrect if Z is even. Only expressions exactly like this are affected: Those where a constant to a power are subtracted from a variable or an expression. Use parentheses to force correct parsing:  $Y - (2 ** Z)$ .

##### MULT RESPONSE:

- (a) All reported frequencies have minima of 1.
- (b) MULT RESPONSE contains no check for alphanumeric variables.

##### RELIABILITY:

Friedman's test (Option 15) cannot be computed by the default computing method because that code skips scales with zero variance, which occur whenever the data are ranks. When Option 15 is selected, Option 14 should be selected as well, to force the alternative computing method.

REPORT:

REPORT loses dummy or string variables in SUMMARY statements.

SAVE ARCHIVE/LIST ARCHINFO:

SAVE ARCHIVE or LIST ARCHINFO will fail if DOCUMENT precedes GET ARCHIVE in the run.

VALUE LABELS:

If a variable name is omitted on the VALUE LABELS card, the value labels of the missing variable destroy the value labels of the previously defined variable instead of generating a syntax error.

WRITE FILEINFO/LIST FILEINFO:

WRITE FILEINFO and LIST FILEINFO produce incorrect N OF CASES values if they are the first procedure and if N OF CASES is UNKNOWN.

```

* * * * *
*           H P - 2 0 0 0           *
* * * * *

```

## \*\*\*\*SPECIAL ANNOUNCEMENT TO HP USERS\*\*\*\*

It is once again time to clean-up the HP-discs by deleting unused programs/files. On Saturday, June 14, 1980, any program/file that has not been referenced\* since January 12, 1980, will be automatically deleted from the system. A log of all deleted items will be kept. A File/Program Reload Request Form (available at the Computing Center) may be used if it becomes necessary to restore a file/program to your library. If you have files/programs in your library that are to be used in the future and you are not sure of the date they were last used, you must access them prior to June 14, 1980, to avoid submitting the reload request form.

\*A program has been referenced when created, read, listed, or executed. A file has been referenced when created, read, written to, or listed.

SUMMER BACKUP SCHEDULE

Daily backup of the HP will be done between 8 a.m. and 8:20 a.m., Monday through Friday. A weekly backup will be performed every Friday between the hours of 8 p.m. and 9:30 p.m. Scheduled shutdowns may occur on weekends due to software development being carried on by the Music Department. At least 48 hour advanced notice of scheduled downtime for preventive maintenance and weekend software development will be provided via system log-on messages.

MINI-SEMINARS

This summer there will be a series of single session mini-seminars for HP-2000 users. These seminars are free of charge and are open to any Hewlett-Packard users. Please make reservations by contacting the Computing Center at 788-2324. Enrollment for each class is limited to 12 people. The seminar schedule is as follows:

(1) Getting Started on the HP-2000 (1 hour)

This one hour seminar is for new users. Topics include: How to log on and log off the system; how to create and purge files; how to enter and edit programs; how to test, run, and save programs; elementary system commands.

A new session begins every Tuesday (June 17 - August 12) in ISB 224, 4-5 p.m.

(2) Introduction to HP-2000 BASIC (2 hour)

This two hour seminar covers fundamentals of the BASIC programming language as implemented on the HP-2000. Statement types, built-in functions, and character string manipulation will be emphasized.

Each session will meet in ISB 224 from 3-5 p.m. Interested individuals have the following dates to choose from: Monday, June 23; Monday, July 7; Monday, July 21; Monday, August 4.

(3) File Processing on the HP-2000 (2 hour)

This two hour seminar will cover the use of BASIC and ASCII disc files in HP-2000 BASIC programs. Included are methods for handling files, and using both sequential access and direct access. This seminar is aimed at intermediate/advanced users.

Each session will meet in ISB 224 from 3-5 p.m. Users may attend on any one of the following dates: Thursday, June 19; Thursday, July 3; Thursday, July 17; Thursday, July 31; Thursday, August 14.



# BEHAVIORAL AND SOCIAL SCIENCE DATA ANALYSIS USING OSIRIS

By Dr. C. Neal Tate, Department of Political Science.

OSIRIS III is one of three major contemporary behavioral and social science data analysis program packages implemented on NTSU's AS 5000 computer system. It is in fact the oldest of the three and is, in my view, considerably more powerful than the best known package, SPSS. Although it appears to be less powerful in some respects than the third of these packages, the newly-acquired SAS, it nevertheless provides the behavioral and social science data analyst with many data management, variable transformation and recoding, and statistical analysis capabilities which are not available in SAS. Hence it is a data analysis system well worth learning -- even by those whose package of first preference may be the more popular SPSS or the more modern SAS.

OSIRIS (Organized Set of Integrated Routines for Investigation with Statistics) was created and is distributed by the Institute for Social Research (ISR) at the University of Michigan. NTSU runs release 2, the final version of OSIRIS III. An important feature of OSIRIS III is that it may also be used to access SRCLIB, another set of programs distributed by ISR, which contains several important analysis programs. The essential documentation for OSIRIS is:

Institute for Social Research (1973). OSIRIS III. Volume 1. SYSTEM AND PROGRAM DESCRIPTION. Release 2 Edition. Ann Arbor: University of Michigan. 846 loose-leaf pages.

There are in addition, five other volumes of documentation for OSIRIS which may be of interest to some users some of the time. In brief, these include:

- Volume 2, ERROR MESSAGES
- Volume 3, Release 2 Edition, SUMMARY OF CONTROL CARDS
- Volume 4, SAMPLE JOBS
- Volume 5, FORMULAS AND STATISTICAL REFERENCES
- Volume 6, PRIMER.

All these are available in the program manuals rack at the Computing Center. In addition, Volume 1 is usually available at the book stores.

OSIRIS, SAS, and SPSS all run in batch mode on the AS 5000. Although there are interactive versions of OSIRIS and SPSS available, there are no current plans to implement them because they lack the full capabilities of the batch versions and apparently require operating systems not planned for the AS 5000. However, it will be possible, with a little practice, to set up and execute these packages by "remote batch" from a terminal, once the MUSIC interactive operating system is implemented on the AS 5000. A user should then be able to set up, execute, and check the results of execution (including printout produced) from his or her terminal before having the results printed on a line printer, or, if the amount of printout is small, at a hard copy terminal. This arrangement should greatly reduce the need to make the familiar trek from office to Computing Center and back which is so dear to the hearts of users of the late IBM 360 and of the current AS 5000 system.

## What OSIRIS Can Do

Many of the things which OSIRIS can do also can be done by SAS and/or SPSS, sometimes more elegantly than OSIRIS can do them. On the other hand, there are some OSIRIS capabilities which are unique, or which OSIRIS performs in a clearly superior manner.

General Data Handling Capabilities. OSIRIS can: copy at one time up to 99 files from any input to any output media, including cards; can sort, merge, subset and edit any data (SAS performs the sort function in a less cumbersome manner for SAS data sets only); process, decode, and produce a new data set from multiply-punched data (SAS and SPSS do not handle multiply-punched data of the kind which was once common in the social sciences in a straightforward manner); check deck order (deck "merge") on any data and pad for missing decks to achieve a uniform rectangular data set (not performed by SAS or SPSS although each requires a complete rectangular data set); produce univariate frequencies for up to four digit variables in raw data for any combination of the codes 0-9, &, - (raw data handling not available in SAS and SPSS).

Building OSIRIS Datasets. OSIRIS, like SAS and SPSS, does most of its data handling and analysis work on specially-documented files which it creates to save the user the necessity of describing the characteristics of his or her data each time it must be manipulated or analyzed. OSIRIS builds these files (called OSIRIS datasets) through its FBUILD program or through a set of cards describing variables called "Type 3 Dictionary Cards." (SAS data sets and SPSS system files may be a bit more thoroughly and flexibly documented than OSIRIS datasets because one is allowed to specify labels for the values of categorical variables, to use mnemonic variable names, if desired, and to utilize slightly longer extended

labels for variable identification names of labels.) OSIRIS will also build elaborately-documented codebooks for any OSIRIS dataset, at the user's option, and has full capabilities for correcting OSIRIS datasets and the "dictionary" files which describe them, once they have been created.

Listing and Manipulating OSIRIS Datasets. OSIRIS can: produce a neatly-organized listing of all or any portion of a dataset (SAS and SPSS have more elegant capabilities in their "report-writing" modes); combine two or more datasets containing the same or an overlapping set of observations but different sets of variables ("match-merge"); create a temporary or permanent subset of any dataset by observations and/or by variables or of any OSIRIS matrix of variables (such as a correlation matrix) by variable; "unbuild" an OSIRIS dataset to "card image" data; lag or lead any variables for time series analysis.

Data Simulation. OSIRIS can create an artificial dataset to match sets of user-supplied functions underlying variables. This capability (unique to OSIRIS) is very useful for teaching applications and Monte Carlo simulations.

Transforming Data. OSIRIS can either temporarily or permanently transform any variable in almost any imaginable way through either of two separate recoding programs (ICON and RECODE). (While it is possible to do in SPSS everything one can do with the OSIRIS recoding and variable transformation programs, it is often much more laborious to do so because of the much greater variety of commands and transformations available in OSIRIS.) In addition, OSIRIS performs aggregations across cases to produce means, sums, standard deviations, counts, etc., for subgroups of the data defined in any of several ways. It has both a simple and a powerful, more complicated program to do this.

Statistical Analysis. OSIRIS will do: univariate frequencies and summary statistics; n-way cross-tabulations accompanied by a large family of measures of association and significance appropriate to tables (SPSS has easier setup of tables, better printout, if value labels are used, and more statistics); matrices of ordinal and categorical correlation for a large number of statistics for 2 x 2 and 3 x 3 tables; scatterplots for pairs of variables; matrices of Pearsonian Correlation Coefficients (in three ways) and of n-order partial correlation coefficients; simple or multiple linear regression including dummy or categorical variables either stepwise or directly, and output residuals for subsequent analyses (SAS is better); multiple classification analysis (a variation of categorical regression -- essentially unique to OSIRIS); two-stage linear regression (SAS is better); one-way and multivariate analysis of variance; Factor Analysis and factor scoring (SAS and SPSS are better); cluster and hierarchical cluster analysis; multidimensional scaling using the Sheppard-Kruskal program (unique to OSIRIS); factor and other geometric configurations comparison, rotation, and analysis (essentially unique to OSIRIS); scaling and scoring; AID or "tree analysis" (unique to OSIRIS); THAID analysis, i.e., "tree analysis" for categorical dependent variables; MNA or Multivariate Nominal Analysis (multiple categorical regression with categorical dependent variable (unique to OSIRIS)).

In addition, programs in SRCLIB, accessible through OSIRIS, perform the following analyses: computation of sampling errors for complex designs; calculation of Cronbach's alpha measuring reliability of multi-term scales; linear or logistic maximum likelihood regression; binary ranking comparisons for N objects ranked by k judges using different numbers of objects and of ranks; Kuder-Richardson formula 3 reliability analysis for internal consistency of Likert scales; Life Table analysis; multidimensional scaling using the Guttman-Lingoes program; fast computation of rank order statistics, univariate and bivariate frequencies; calculate point biserial correlation coefficients; compute Bohrnstedt Reliability Validity checks from correlation and standard deviation matrices; fast computation of univariate statistics for large numbers of cases.

Two classical multivariate techniques, discriminant and canonical correlation analysis, are not directly available in OSIRIS, although they are in both SAS and SPSS.

Finally, it is worth noting that OSIRIS datasets can be read directly by either SAS or SPSS so that information contained in the OSIRIS dictionary documenting the OSIRIS dataset can be utilized to simplify analyses in SAS or SPSS. This means that, with very little effort, analysis techniques which are available in SAS or SPSS can be applied to data which are in OSIRIS format. Unfortunately, the reverse is not true. OSIRIS III does not have the capability to read SAS or SPSS files directly, and is thus unable to utilize the documenting information they contain. Given this one-way compatibility between OSIRIS, SAS and SPSS, the wise data analyst would build his or her basic data file through OSIRIS, thereby making all three packages' capabilities available with no necessity for redocumenting his file when shifting from one to another package.

If you are interested in learning more about OSIRIS III, please drop me a note saying so. I have in the past organized informal OSIRIS seminars for interested persons. It might be possible to do so again in late summer or early fall.



JCL For Executing OSIRIS

The basic EXEC statement is:

```
// EXEC OSIRIS,REGION=110K
```

The default region for OSIRIS is currently 90K; this works for some programs, but is a bit tight for most analysis programs. Hence the REGION parameter.

To access a program in SRCLIB through OSIRIS, it is first necessary to mount the disk on which SRCLIB currently resides. Thus the necessary JCL is:

```
// EXEC ICPRO1
// EXEC OSIRIS,REGION=110K,LIB1='USER1.D0411.P1113.SRCLIB'
```

Other necessary JCL and OSIRIS JCL conventions are described in the OSIRIS manual.

INDEX TO PAST ISSUES

In order to utilize BENCHMARKS to its fullest capacity and avoid redundancies, an index of all previous issues will henceforth be included in each new publication. Individuals can then use this as a guide to refer back to articles on topics that pertain to their interests.

| <u>SUBJECT</u>                                                                                    | <u>VOL/NO</u> | <u>MONTH/YEAR</u> | <u>PAGE</u> |
|---------------------------------------------------------------------------------------------------|---------------|-------------------|-------------|
| 1. CHARGE RATES<br>Charge Rates                                                                   | 1/1           | Jan/1980          | 1           |
| 2. GAINING ACCESS TO COMPUTING FACILITIES<br>So You Want To Compute                               | 1/1           | Jan/1980          | 2           |
| 3. HARDWARE<br>Hardware - What and Where                                                          | 1/1           | Jan/1980          | 3           |
| 4. JOB CLASSES<br>Job Class Selection                                                             | 1/1           | Jan/1980          | 5           |
| 5. SERVICES AVAILABLE TO USERS<br>Services Available to Users of the NTSU Computing<br>Facilities | 1/1           | Jan/1980          | 2           |
| 6. SOFTWARE<br>Software - What We Have and What We Support                                        | 1/1           | Jan/1980          | 4           |
| 7. STATISTICAL PACKAGES<br>Attention SPSS Users                                                   | 1/1           | Jan/1980          | 7           |

FUTURE SEMINARS: USER FEEDBACK REQUESTED

The Computing Center is planning to begin offering seminars on topics of interest to users of our facilities. Some suggested topics have been: (A) SPSS (B) Utilities (C) JCL (D) SAS (E) HP (Note: Some of these are already scheduled for this summer - see this issue.) (F) Setting up research projects for computer analysis.

Please contact us, either at the address below, by telephone (788-2324), or by coming by the Computing Center, and let us know your reaction to the idea of seminars. Please specify the topic(s) you would like a seminar in and any other information you feel would be helpful along the same line.

\*\*\*\*\*  
 \*  
 \* BENCHMARKS is a quarterly publication of the NTSU Computing Center. Reader/user feedback is \*  
 \* encouraged, so send all letters, suggestions, etc., to: NTSU Computing Center \*  
 \* NT Box 13495 \*  
 \* Denton, TX 76203 \*  
 \*  
 \* Claudia Putnam . . . . . BENCHMARKS Editor \*  
 \*  
 \* Richard Harris . . . . . Director of Computer Systems \*  
 \*  
 \* Douglas Lillard. . . . . Assistant Director, \*  
 \* Academic Computing \*  
 \* \*\*\*\*\*

*BENCHMARKS is the Computing Center newsletter aimed at the NTSU academic community. A benchmark is a point of reference from which measurements of any sort are made, and that name was chosen to symbolize the intent of this publication - to provide a point of reference for the users of the North Texas State University computing facilities. BENCHMARKS was created to provide a channel of communications between the NTSU Computing Center and its users - a channel, hopefully, that will flow both ways, enabling both the users and the Computing Center staff to gain new insights into the ever changing world of academic computing.*